# README File for 2010 Census Congressional District Summary File (116th Congress) Delivered via FTP

---

*Note: Users processing these FTP files in a Windows environment should read carefully the File Information section of this document.*

---

**Contents**

**About the FTP Application**

**This FTP (File Transfer Protocol) application is intended for experienced users of census data, zip files, and spreadsheet/database software.** It provides quick access for data users such as State Data Centers and news media needing to begin their analysis immediately upon data release.

**The FTP Directory**

The directory is http://www2.census.gov/census_2010/08-SF1_Congressional_Districts_116/. When the 2010 Census Congressional District Summary File (116th Congress) (CD116) data file is added to each state's directory, it will contain a single zip package with the geoheader file and 48 data segments. See below for more information on Segmented Data.

Users of the FTP application need to unzip the package after downloading, then import the data into the spreadsheet/database software of their choice for data analysis and table presentation. *We are unable to provide one-on-one support for applications of the data to specific spreadsheets or data base software.*

**File Information**

The "zip" archive files are in the Unicode Transformation Format (UTF-8). The geographic header file contains fixed fields, while the data segments are in a comma-delimited format. Both the geographic header and the data files contain geographic linking fields. These files were created in a Linux environment, but should be usable from any operating system. Standard software packages can be used to import and manipulate these data files. These are text files,

however, the file extension is **not** '.txt'. The user will need to rename the files with a .txt extension for import into some software packages.

The resulting file will meet the ANSI MS-DOS/Windows standard. If the data are being processed in a LINUX environment, they can be unzipped using any standard Linux ZIP/UnZIP package.

**Note to Users:**

Modifications to field types* are required when importing the Geographic Header Record file and File01 through File48 into MS Windows-based programs:

➢ Fields classified as numeric (N) should be imported as long integers unless the field has been described as having decimals. These fields should be imported as double.

➢ AREALAND and AREAWATR should be imported as text.

➢ Failure to make these changes may result in missing data upon import.

*For additional information about modifying field types see, https://www2.census.gov/programs-surveys/decennial/rdo/technical-documentation/CD116_TechnicalDocumentation.pdf.

**Segmented Data**

The data in the 2010 CD116 and other 2010 Census summary files are segmented. This is done to manage the volume of data and to facilitate exporting into spreadsheet or database software. The data and the corresponding geographic information for an individual state are known as the file set. Because of the large size of the tables, the CD116 file set will be broken into 49 files: a geographic header record file and 48 data segment files. To get a complete set of the 2010 CD116 data, users must download the geographic header file and all the data file segments in the package.

Table 1 below provides the file/table details. When using Table 1 you must take into consideration that the first 5 fields of the data file contain identification information (FILEID, STUSAB, CHARITER, CIFSN, and LOGRECNO) which makes some segments have more than 255 fields. These segments are in bold font in Table 1. Some manipulation of the files will be required for software with a field limit of 255. Also, due to new disclosure avoidance requirements, this product is omitting any of the tables previously only produced at the tract or county level. This results in many of the data segments containing no data, represented by italicized text in Table 1. In order to preserve the structure of the data segments to align with the original Summary File 1 product and previously created code to work with these data segments, they were produced and included despite not containing data.

**Table 1.  File/Table Segmentation**

| Data file segment number | Number of data cells | Starting table (matrix) number | Ending table (matrix) number |
| --- | --- | --- | --- |
| 1 | 1 | P1 | P1 |
| 2 | 6 | P2 | P2 |
| 3 | 194 | P3 | P9 |
| 4 | 239 | P10 | P14 |
| 5 | 245 | P15 | P30 |
| **6** | **254** | **P31** | **P49** |
| **7** | **251** | **P50** | **P12E** |
| **8** | **253** | **P12F** | **P17A** |
| 9 | 249 | P17B | P28I |
| **10** | **252** | **P29A** | **P29I** |
| **11** | **254** | **P31A** | **P34E** |
| **12** | **251** | **P34F** | **P38E** |
| 13 | 240 | P38F | P39H |
| 14 | 20 | P39I | P39I |
| *15* | *251* | *PCT1* | *PCT8* |
| *16* | *59* | *PCT9* | *PCT11* |
| *17* | *209* | *PCT12* | *PCT12* |
| *18* | *188* | *PCT13* | *PCT20* |
| *19* | *216* | *PCT21* | *PCT22* |
| *20* | *209* | *PCT12A* | *PCT12A* |
| *21* | *209* | *PCT12B* | *PCT12B* |
| *22* | *209* | *PCT12C* | *PCT12C* |
| *23* | *209* | *PCT12D* | *PCT12D* |
| *24* | *209* | *PCT12E* | *PCT12E* |
| *25* | *209* | *PCT12F* | *PCT12F* |
| *26* | *209* | *PCT12G* | *PCT12G* |
| *27* | *209* | *PCT12H* | *PCT12H* |
| *28* | *209* | *PCT12I* | *PCT12I* |
| *29* | *209* | *PCT12J* | *PCT12J* |
| *30* | *209* | *PCT12K* | *PCT12K* |
| *31* | *209* | *PCT12L* | *PCT12L* |
| *32* | *209* | *PCT12M* | *PCT12M* |
| *33* | *209* | *PCT12N* | *PCT12N* |
| *34* | *209* | *PCT12O* | *PCT12O* |
| *35* | *245* | *PCT13A* | *PCT13E* |
| *36* | *245* | *PCT13F* | *PCT19B* |
| *37* | *237* | *PCT19C* | *PCT20E* |
| ***38*** | ***254*** | ***PCT20F*** | ***PCT22F*** |
| *39* | *63* | *PCT22G* | *PCT22I* |
| *40* | *234* | *PCO1* | *PCO6* |
| *41* | *156* | *PCO7* | *PCO10* |
| 42 | 1 | H1 | H1 |
| 43 | 6 | H2 | H2 |
| 44 | 249 | H3 | H11F |
| **45** | **255** | **H11G** | **H17C** |
| 46 | 126 | H17D | H17I |
| *47* | *74* | *HCT1* | *HCT4* |
| *48* | *47* | *PCT23* | *PCT24* |

The explanation below for linking the two data files requires specific location information about the geographic header. These are located in Chapter 2, *How to Use This Product*, of the Technical Documentation, https://www2.census.gov/programs-surveys/decennial/rdo/technical-documentation/CD116_TechnicalDocumentation.pdf.

A unique logical record number (LOGRECNO) in the geographic header is assigned to all files for a specific geographic entity; all records for that entity can be linked together across files. The geographic header **record layout** is identical across all electronic data products from the 2010 Census. However, the content is product specific. Some header fields that appear in both file types (geographic header and fileXX) are not used. For example, the characteristic iteration (CHARITER) field is filled in the 2010 Census Summary File 2 data product, but in the 2010 CD116, it is always coded as 000.

**File Record Layout**

For a layout of the data tables, see https://www2.census.gov/programs-surveys/decennial/rdo/technical-documentation/CD116_TechnicalDocumentation.pdf, and select Chapter 6, *Data Dictionary*.